An Efficient Approach for DNA Fractal-based Image Encryption

Qiang Zhang\textsuperscript{1}, Shihua Zhou\textsuperscript{2} and Xiaopeng Wei\textsuperscript{1,2}
\textsuperscript{1}Key Laboratory of Advanced Design and Intelligent Computing (Dalian University) Ministry of Education, Dalian, 116622, China
\textit{Email Address: zhangq@dlu.edu.cn}
\textsuperscript{2}School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China
\textit{Email Address: shihuajo@gmail.com, xpwei@dlu.edu.cn}

Received June 22, 200x; Revised March 21, 200x

Security of the image information has become more and more important. At present, researchers have paid much attention to DNA cryptography-based image encryption. In this paper, an efficient encryption approach is proposed, in which, we do not use DNA biological operation to implement image encryption, but DNA sequences are used as the secret keys. The permutation process is implemented by using Hao’s fractal sequence representation and the diffusion process is used to alter the gray values. According to the simulation experiment and performance analysis, this approach is feasible and effective.

\textbf{Keywords:} permutation, DNA fractal, image encryption.

1 Introduction

owing to the widespread transmission over various communication networks, security of the image information has become more and more important [1-3]. It has been noticed that the traditional text encryption approaches fail to protect the image information effectively due to some special properties of the image and some specific requirements of the image processing, such as enormous size, strong redundancy of uncompressed data and the high correlation coefficient [4, 5]. Therefore, the research of the image encryption approach has become a important research field since the early 1990s so that the digital images are protected from illegal acquisition, copy and modification.

At present, DNA cryptography-based image encryption has become a hot research field. DNA cryptography [6, 7] is a new born cryptography, in which DNA is used as information carrier and the modern biological technology is used as implementation tool, and the vast parallelism, exceptional energy efficiency and extraordinary information density inherent
in DNA molecules are explored for cryptographic purposes such as encryption, authentication, signature, and so on. The main security basis depends on the restriction of biotechnology, which has nothing to do with computing power. For example, Clelland et al. [8] proposed an approach based on micro-dots. In this approach, the researchers produced artificial DNA strands, which contained secret messages. A triplet encodes one character or number. It is a simple substitution cipher which encodes characters into DNA sequences. Leier et al. [9] encoded binary information into DNA sequences. A short DNA sequence represents the binary 1, another one represents 0. They bind directly to the corresponding binary information. Wong et al. [10] presented a new approach, which is able to store data in living organisms. The data are translated into a DNA sequence which is inserted into a vector.

An image encryption approach based on DNA fractal is presented in this paper. This approach is not the one based on the real DNA cryptography, but uses the natural DNA sequences as the secret keys. The main security basis depends on the variety of DNA sequence, since the natural DNA sequence is the natural one-time pad. The permutation process is implemented by using Hao’s fractal sequence representation [11] in particular. This approach is feasible and effective according to the simulation experiment and the performance analysis.

2 The Introduction of the Permutation Approach

2.1 Hao’s Fractal Sequence Representation

Hao et al.[11] proposed a DNA fractal sequence representation approach, in which, given a complete genome of length \( N \), i.e., a linear or circular DNA sequence made of \( N \) letters from the alphabet \( \sum = \{A, C, G, T\} \). He defines a mapping that maps the four letters \( G, C, A \) and \( T \) to the base 4 number system:

\[
\alpha : \{G, C, A, T\} \to \{0, 1, 2, 3\}.
\] (2.1)

The coordinates for the counter of the first \( K \)--string are

\[
x = \sum_{i=1}^{K} 2^{k-i} [\alpha(s_i) \gg 1].
\] (2.2)

\[
y = \sum_{i=1}^{K} 2^{k-i} [\alpha(s_i) \& E].
\] (2.3)

where \( \gg 1 \) means the bitwise operation "right-shift by one bit", \( \& \) is the bitwise operator "logical and" and \( E \) is the binary number 1.
2.2 Generation of the modified secret key

Here, we introduce the modified secret key. The purpose of this modified secret key is to modify the key of the above permutation approach so that the permutation approach becomes more secure. The gray image $A_{m \times n}$ is $m \times n$ in size. $a_{ij}$ is the gray value of the image pixel, where $i = 1, 2, ..., m$, $j = 1, 2, ..., n$. Firstly, we use Eq.(2.4) and Eq.(2.5) to calculate $k_1$, $k_2$.

\[
k_1 = \sum_{i=1}^{\lfloor \frac{m}{2} \rfloor} \sum_{j=1}^{n} a_{ij}.
\]

(2.4)

\[
k_2 = \sum_{i=\lfloor \frac{m}{2} \rfloor+1}^{m} \sum_{j=1}^{n} a_{ij}.
\]

(2.5)

As is well known, the range of a pixel gray value is 0-255, so the range of $k_1$ is $0 - (\lfloor \frac{m}{2} \rfloor \times n) \times 255$ and the range of $k_2$ is $0 - ((m - \lfloor \frac{m}{2} \rfloor) \times n) \times 255$. Secondly, we convert $k_1$ and $k_2$ into two binary sequences $B_{k1}$ and $B_{k2}$. The length of $B_{k1}$ is $l_1$, and the length of $B_{k2}$ is $l_2$. The permutation key sequence is converted into the binary sequence $B_k$. The length of $B_k$ is $l_k$. $B_{k1}$ and $B_{k2}$ are enlarged to $l_1 = l_k$ and $l_2 = l_k$. We put $B_{k1}$ and $B_{k2}$ into reverse order and $B_{k1}^1$ and $B_{k2}^1$ are gained. The purpose is that the permutation key is sensitive to the slight change of the gray value. At last, $B_k$ and $B_{k1}^1$, $B_{k2}^1$ are XORed bit-by-bit in turn respectively, and $B_m$ is gained. Then we convert $B_m$ into DNA sequence $D_m$.

2.3 A Simple Example

Through a simple example, the permutation approach is introduced in detail. For the original gray image (Fig.2.1 (a)) that is $3 \times 3$ in size, we list the gray value matrix (Fig.2.1 (b)). Then, we calculate $k_1$ and $k_2$, and $k_1 = 510$ and $k_2 = 523$. we convert $k_1$ and $k_2$ into two binary sequences $B_{k1}$ and $B_{k2}$. The length of $B_{k1}$ is $l_1$, and the length of $B_{k2}$ is $l_2$. The permutation key sequence is converted into the binary sequence $B_k$. The length of $B_k$ is $l_k$. $B_{k1}$ and $B_{k2}$ are enlarged to $l_1 = l_k$ and $l_2 = l_k$. We put $B_{k1}$ and $B_{k2}$ into reverse order and $B_{k1}^1$ and $B_{k2}^1$ are gained. The purpose is that the permutation key is sensitive to the slight change of the gray value. At last, $B_k$ and $B_{k1}^1$, $B_{k2}^1$ are XORed bit-by-bit in turn respectively, and $B_m$ is gained. Then we convert $B_m$ into DNA sequence $D_m$.

So the secret key sequence becomes $\{ATGCATCACTACAGTACACGCGTCACATACGCGTCACATACGCGTCACATACGCGTCACATACGCGTCACAT\}$ (Fig.2.1 (d)). According to 2.1, we gain the DNA fractal image (Fig.2.1 (e)) and the permutation matrix.
is gained (Fig.2.1 (f)). Then, the original gray matrix is rearranged in the light of the above matrix, and the new gray value matrix (Fig.2.1 (g)) is gained. At last, we gain the permutation image (Fig.2.1 (h)) on the basis of the new gray value matrix.

3 The Introduction of the Diffusion Approach

3.1 Addition and Subtraction Algebraic Operation for DNA sequence

Based on the DNA sequence a lot of biology operations and algebraic operations have been presented by researchers [12], such as the addition operation, the subtraction operation, XOR and so on. The addition and subtraction operations are performed in the light of the operation is introduced in Ref. [13]. The details of the addition and subtraction rules are shown in Table 3.1 and 3.2. In this paper, we will use these wonderful addition rules to diffuse the pixel values of the original image.
An Efficient Approach for DNA Fractal-based Image Encryption

3.2 The New Definition of Multiplication Operation

In this paper, we define a new rule of DNA matrix multiplication operation. Two DNA sequences \(d_1\) and \(d_2\) are given, whose lengths are \(l_{d1}\) and \(l_{d2}\). Reconstruct \(d_1\) and \(d_2\) as two matrices \(D_1(l_{d1}, 1)\) and \(D_2(1, l_{d2})\). Performing the multiplication operation for \(D_1(l_{d1}, 1)\) and \(D_2(1, l_{d2})\), we obtain the matrix \(D\) whose size is \(l_{d1} \times l_{d2}\). The formulas is as follow:

\[
D_1 \times D_2 = \begin{bmatrix}
  d_{11}d_{21} & d_{11}d_{22} & \cdots & d_{11}d_{2l_{d2}} \\
  d_{12}d_{21} & d_{12}d_{22} & \cdots & d_{12}d_{2l_{d2}} \\
    \vdots & \vdots & \ddots & \vdots \\
  d_{1l_{d1}}d_{21} & d_{1l_{d1}}d_{22} & \cdots & d_{1l_{d1}}d_{2l_{d2}} 
\end{bmatrix} \tag{3.1}
\]

And the rule of \(d_1 \times d_2\) is shown in Table 3.3.

<table>
<thead>
<tr>
<th>×</th>
<th>T</th>
<th>A</th>
<th>C</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>A</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>G</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>T</td>
</tr>
</tbody>
</table>

For example, we set \(D_1 = (ATCG)^{-1}\) and \(D_1 = (GATC)\).

\[
D_1 \times D_2 = \begin{bmatrix}
  AG & AA & AT & AC \\
  TG & TA & TT & TC \\
  CG & CA & CT & CC \\
  GG & GA & GT & GC 
\end{bmatrix} \begin{bmatrix}
  C & T & G & A \\
  G & A & T & C \\
  A & G & C & T \\
  T & C & A & G 
\end{bmatrix} = \begin{bmatrix}
  C & T & G & A \\
  G & A & T & C \\
  A & G & C & T \\
  T & C & A & G 
\end{bmatrix} \tag{3.2}
\]
4 The Flow of the Encryption Algorithm

In this paper, the original image is confused in the light of the permutation sequence that is generated by using the natural DNA sequence modified and Hao’s fractal sequence representation. Fig.4.2 is the flow chart of this algorithm. The encryption approach is as follow:

Step1: Input the original image $A_0$ is $m \times n$ in size, where $m$ and $n$ are rows and columns of the image respectively.

Step2: Convert the image into a binary matrix, then carry out DNA encoding for the binary matrix, to obtain a DNA matrix $D_0$, the size of $D_0$ is $m \times (n \times 4)$.

Step3: Divide $D_0$ into some equal blocks $D_0\{i, j\}, i = 1, 2, ..., m, j = 1, 2, ..., n$, where the size of blocks is $1 \times 4$. According to section 2.2, the modified secret keys are generated. The natural DNA sequence is modified by the modified secret keys is regarded as the secret key sequence $k_s$. Generate two permutation sequences $X = \{x_1, x_2, ..., x_m\}$, $Y = \{y_1, y_2, ..., y_n\}$ by Hao’s fractal and the secret key sequence $k_s$. Let the location value of sequences $X, Y$ be row coordinates and column coordinates of $D_0\{i, j\}$. The scrambling DNA matrix $D_1$ is gained.

Step4: Two DNA sequences $k_1$ and $k_2$ are input, whose lengths are $m$ and $n \times 4$. Reconstruct $k_1$ and $k_2$ as two matrices $K_1(m, 1)$ and $K_2(1, n \times 4)$. Performing the multiplication operation for $K_1(m, 1)$ and $K_2(1, n \times 4)$, we obtain the matrix $D_{\text{template}}$ whose size is $m \times (n \times 4)$ in the light of the rules of section 3.2.

Step5: Add $D_1$ and $D_{\text{template}}$ according to the rules in section 3.1, obtaining the result as $D_2$.

Step6: A chaotic sequence $z_1$ is produced by Logistic Map, whose length is $m \times n \times 4$. We convert $z_1$ into one matrix $Z$ whose size is $m \times (n \times 4)$. Then we use the following threshold function $f(x)$ to get a matrix $Z^1$:

$$f(x) = \begin{cases} 0 & 0 < z(i, j) \leq 0.5 \\ 1 & 0.5 < z(i, j) \leq 1 \end{cases} \quad (4.1)$$

Step7: If $Z^1\{i, j\} = 1$, $D_2$ is complemented, otherwise it is unchanged. After the complementing operation, we get a new coding matrix $D_3$.

Step8: Carry out the inverse process of the step 2, for the DNA matrix $D_3$, then we gain the gray value matrix $I(m, n)$. We gain the encrypted image $A_1$.

The processes of the decryption algorithm is the inverse processes of the encryption one. Receivers obtain the encrypted image $A_1$ through the insecure channel and the secret keys through the secure channel from the senders. Receivers use the keys to decrypt the encrypted image in the light of the reverse operation of the encryption algorithm.
5 Experimental Results

In this paper, simulation results are given to demonstrate the validity and the effectiveness of the presented approach. For the original gray image that is $256 \times 256$ in size, we use Matlab7.1 to simulate the algorithm on the condition that we set the keys that are (8, human -Globin, SRY HMG-BOX, 256, 0.6).

Fig.5.3 is the experimental results show the encrypted image and decrypted image. Fig.5.3 (a) shows the original image, and Fig.5.3 (b) is the permutation image. And Fig.5.3 (c) is the encrypted image. When the true keys are obtained, the encrypted image is executed in the light of the decryption algorithm, and we can gain the decrypted image shown in Fig.5.3 (d). From the experimental results, the encryption approach is feasible and satisfactory.

6 Algorithm Performance Analysis

As is well known, a good algorithm should be robust against all kinds of attacks. In this section, the presented algorithm is analyzed by different security measures. They include key space analysis, key sensitivity analysis, correlation analysis, the gray histogram analysis, differential attack, information entropy, and so on. The security analysis demonstrates that this algorithm is feasible and effective.
6.1 Key Space Analysis

We must be assumed that an attacker knows all things except for the secret key in the light of the hypothesis that is proposed by Kerckhoff. Therefore, the key space must be large enough so as to repel the brute force attack. In the nature, DNA sequence is various, and the length also has the considerable difference. In case of the same DNA sequence, since the difference of the length and the initial position, a segment is quite different from others. So the set of the natural DNA sequences is the natural one-time pad. With the fast development of genetic engineering, the scale of gene bank becomes larger and larger. In this encryption approach, we use the natural DNA sequences as the secret keys. So the key space of our approach is large enough to resist brute-force attacks. Moreover, there’s no need for us to worry about key management. We only need transmit the name, location and number of DNA sequences. Receiver can gain DNA sequences by http://www.ncbi.nlm.nih.gov/(The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information).
6.2 Key Sensitivity Analysis

In order to further analyze key sensitivity, we test the algorithm under the worry decryption keys. For example, the decryption keys are (8, SRY HMG-BOX, human -Globin, 256, 0.6). Fig.6.4 (a) shows the original image, and Fig.6.4 (b) is the encrypted image. And Fig.6.4 (c) is the decrypted image under the wrong key. Fig.6.4 (d) is the decrypted image. The decrypted image under the wrong key is different from the original image, and we barely distinguish any information of the original image.

![Figure 6.4: The Experimental Results of the Wrong Key](image)

6.3 Permutation Extent Analysis

We use Fig.6.5 (a) that is $256 \times 256$ in size and the color is pure write besides the solid black block in the middle to simulate the permutation process in order to analyze permutation extent. Fig.6.5 (b) shows the experimental result.

From Fig.6.5 (b), we can see that the solid black block in the middle is evenly spread to the whole image, and it is proved that the new image permutation method has better permutation effect.
6.4 Correlation Coefficient Analysis

In this section, a statistical test on the correlation between two adjacent pixels of the encryption image has been carried out. From the horizontal, vertical and diagonal direction, 3000 pairs of adjacent pixels are randomly selected. The correlation coefficients of the pixel pairs are calculated by using the following formulas.

\[ E(x) = \frac{1}{N} \sum_{i=1}^{N} x_i. \]  

(6.1)

\[ D(x) = \frac{1}{N} \sum_{i=1}^{N} (x_i - E(x))^2. \]  

(6.2)

\[ \text{cov}(x, y) = \frac{1}{N} \sum_{i=1}^{N} (x_i - E(x))(y_i - E(y)). \]  

(6.3)

\[ \gamma(x, y) = \frac{\text{cov}(x, y)}{\sqrt{D(x)} \times \sqrt{D(y)}}. \]  

(6.4)

The results are listed in Table 6.4. The correlation coefficients of two horizontally adjacent pixels of the images before and after encryption are illustrated in Fig. 6.6(a) and (b), respectively. From the table and the figures, we can see that the correlation coefficient of the adjacent pixels in the original image is very high, and the one of the encryption image is far lower than the one of the original image.
Figure 6.6: Correlation of Two Horizontal Adjacent Pixels

Table 6.4: Correlation coefficients of two adjacent pixels

<table>
<thead>
<tr>
<th>Direction</th>
<th>The original image</th>
<th>The encrypted image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal</td>
<td>0.9460</td>
<td>0.0099</td>
</tr>
<tr>
<td>Vertical</td>
<td>0.9218</td>
<td>-0.0264</td>
</tr>
<tr>
<td>Diagonal</td>
<td>0.8776</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

6.5 The Gray Histogram Analysis

Considering the statistical analysis, Fig.6.7(a) and (b) show the gray histograms of the image before and after the encryption, respectively. Comparing two histograms, we find that the pixel gray values of the original image are concentrated on some values, but the histogram of the encrypted image is very uniform.
6.6 Differential Attack

It is called differential attack that attackers often make a slight change to the original image, and use the proposed algorithm to encrypt for the original image before and after changing, through comparing two encrypted image to find out the relationship between the original image and the encrypted image. Here, the encrypted image is called "test1", and the encrypted image after changing the first pixel gray value from "test1" is called "test2". Researchers usually use NPCR (number of pixels change rate) and UACI (unified average changing intensity) as two criterions to examine the performance of resisting differential attack. Here, we use Eq.6.5-6.7 to calculate NPCR and UACI between "test1" and "test2".

\[
C(i, j) = \begin{cases} 
0, & \text{if } T_1(i, j) = T_2(i, j) \\
1, & \text{if } T_1(i, j) \neq T_2(i, j)
\end{cases}
\]  

(6.5)

\[
\text{NPCR} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} C(i, j)}{M \times N} \times 100\%.
\]

(6.6)

\[
\text{UACI} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} |T_1(i, j) - T_2(i, j)|}{255 \times M \times N} \times 100\%.
\]

(6.7)

where \( M \) and \( N \) are the height and width of the image, \( T_1(i, j) \) and \( T_2(i, j) \) denote the pixel value of "test1" and "test2" on the location \((i, j)\). We obtain the result, NPCR=99.62\% and UACI=33.36\%, from the simulation. This result demonstrates that our algorithm has a strong ability to resist differential attack.

6.7 Information Entropy

The information entropy is defined as expressing the degree of uncertainty [14]. We can also use it to express uncertainty of the image information. The information entropy can measure the distribution of gray values in the image, and the value of the information entropy is 8. An effective encryption algorithm should make the information entropy tend to 8. We obtained an information entropy \( H = 7.9936 \), that is very close to 8. It is can be seen that the proposed algorithm is very effective.

6.8 Comparison with Chaos-based Algorithms

A large number of chaos-based image encryption schemes have been proposed. Unfortunately, many of these schemes have been found insecure, especially against known and/or chosen-plaintext attacks. At present, a lot of research on cryptanalysis of chaos-based image encryption scheme is been proposed [15, 16]. However, it is invalid to break
our algorithm by the above cryptanalysis. Moreover, the researchers [6, 17] have found that the traditional image encryption schemes except for the one-time pad own only computational security. The main keys of our algorithm are the natural DNA sequences. Since the set of the natural DNA sequences is the natural one-time pad. What’s more, with the fast development of genetic engineering, the scale of gene bank becomes larger and larger. Thus, our algorithm is more secure. To make a comparison, the performance of our algorithm is as better as the one of chaos-based algorithms [18]. So, our algorithm is feasible and satisfactory.

7 Conclusion

A new approach that is to use Hao’s Fractal sequence representation to finish the permutation process for image encryption is proposed in this paper. Our approach uses the natural DNA sequence is modified by the modified keys as the secret key of the permutation process. The diffusion process is used to alter the gray values. Experimental result shows that this approach is simple and feasible. From performance analysis, our approach meets the corresponding security level and can effectively resist exhaustive attacks.

Acknowledgements

This work is supported by the National High Technology Research and Development Program (“863” Program) of China (No.2009AA01Z416), the National Natural Science Foundation of China (No.30870573), the Program for Liaoning Science and Technology Research in University (No.LS2010179), and by the open fund of Key Laboratory of Advanced Design and Intelligent Computing, (Dalian university) Ministry of Education (No. ADIC2010011). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

Q. Zhang is a professor at Dalian University, Dalian, China. His research interests are intelligent computing and computer animation. He is author of more than 50 articles published in international peer reviewed journals. Now he has served as editorial boards of seven international journals and has edited special issues in journals such as Neurocomputing and International Journal of Computer Applications in Technology.
Shihua Zhou is a Ph.D. student at School of Mechanical Engineering, Dalian University of Technology, Dalian, China. His research fields are DNA computing and image security. In June 2009, she obtained an M.Eng. at Dalian University, Dalian, China.

Xiaopeng Wei is a professor and head of Key Laboratory of Advanced Design and Intelligent Computing (Dalian University), Ministry of Education, Dalian, 116622, China. His research areas include computer animation and intelligent CAD. He has published over 120 papers in international peer reviewed journals. Currently he is at Dalian University, Dalian, China.